



L'approche systèmes complexes de la cognition sociale : une façon de penser l'auto-transformation du social ?

David Chavalarias

► To cite this version:

David Chavalarias. L'approche systèmes complexes de la cognition sociale : une façon de penser l'auto-transformation du social ?. 2006. halshs-00088411

HAL Id: halshs-00088411

<https://shs.hal.science/halshs-00088411>

Preprint submitted on 2 Aug 2006

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

L'approche systèmes complexes de la cognition sociale : une façon de penser l'auto-transformation du social ?

David Chavalarias*

2^{ème} Congrès de l'Association Française de Sociologie
Groupe *ad hoc Sociologie et Systèmes Complexes*
Bordeaux, 5 septembre 2006

Résumé

Certains scientifiques, sociologues ou économistes en particulier, se sont récemment emparés du terme de *cognition sociale* utilisé en psychologie en lui donnant un second sens, à savoir, une cognition distribuée sur l'ensemble des individus composant une société. L'enjeu est alors de savoir " en quoi les aptitudes sociales humaines rendent compte du développement d'une cognition collective complexe ". Nous défendons ici l'idée que la cognition sociale est un raffinement du concept de cognition distribuée plutôt qu'un équivalent. C'est un phénomène propre aux sociétés humaines qui permet de penser leur propriété d'auto-transformation. Nous montrons comment l'imitation permet de penser ce phénomène, comme le suggérait en son temps Gabriel Tarde.

Mots-clef : cognition sociale, auto-organisation, imitation, téléologies individuelles, différenciation sociale.

Introduction

Certains scientifiques, sociologues ou économistes en particulier, se sont récemment emparés du terme de *cognition sociale* utilisé en psychologie¹ en lui donnant un second sens, à savoir, une cognition distribuée sur l'ensemble des individus composant une société [Bourgine(2004)]. L'enjeu est alors de savoir " en quoi les aptitudes sociales humaines rendent compte du développement d'une cognition collective complexe " [Conein(2005)] . Nous défendons ici l'idée que la cognition sociale est un raffinement du concept de cognition distribuée plutôt qu'un équivalent. C'est un phénomène propre aux sociétés humaines. Cette notion doit en effet tenir compte d'un certain nombre de caractéristiques de la cognition humaine, à savoir :

*Centre de Recherche en Epistémologie Appliquée(CREA), Ecole Polytechnique, Paris, France. Contact: david.chavalarias@polytechnique.edu, [Http://www.chavalarias.com](http://www.chavalarias.com)

¹Ensemble d'attitudes permettant la compréhension intuitive des relations sociales.

- reconnaissance de la part du sujet de l'autre comme acteur intentionnel agissant dans et sur son environnement, et prise en compte (par anticipation ou imitation) des actions de celui-ci,
- modification de la constitution même du sujet (ses croyances, ses représentations, ses règles, ses buts, etc.) au cours de ses interactions sociales.

Nous allons tenter ici d'approfondir le concept de cognition sociale ainsi que les apports potentiels d'une approche systèmes complexes en rendant hommage à l'un des précurseurs de la cognition sociale, trop vite oublié de son temps, Gabriel Tarde.

1 Cognition distribuée et cognition sociale

La cognition distribuée peut-être définie comme un traitement de l'information effectué par un ensemble d'entités éventuellement hétérogène. Par exemple, [Hutchins(1995)] défend l'idée d'un traitement de l'information au sein d'un système hétérogène homme/machine. Lorsque que le système est homogène, on emploie souvent le terme de *cognition collective* ou bien d'*intelligence collective*, comme c'est le cas pour l'étude des comportements collectifs des insectes sociaux [Bonabeau and Theraulaz(1994)]. Pour comprendre en quoi le terme de cognition sociale peut se voir attribuer un sens qui est spécifique aux sociétés humaines, donnons un exemple typique de cognition distribuée : la régulation de la température au sein d'une ruche.

Une ruche pour bien fonctionner nécessite une certaine régulation thermique interne. Pour résoudre ce problème, la Nature a trouvé un système ingénieux qui s'appuie sur le traitement distribué de l'information portant sur la température ambiante. Lorsque la température extérieure augmente, certaines abeilles se mettent à battre des ailes de manière à créer une ventilation à l'intérieur de la ruche. Mais les abeilles ne sont pas toutes également sensibles à la température et disposent de seuils différents à partir desquels elles vont se mettre à battre des ailes. Il en résulte un recrutement progressif des abeilles à la tâche de ventilation lorsque la température augmente qui dépend de la distribution des seuils dans la population. Cette distribution de seuils est quant à elle déterminée génétiquement. Ce fonctionnement qui permet une régulation de la température tout en n'immobilisant que le nombre d'abeilles nécessaire est une adaptation de l'espèce à son environnement.

Dans cet exemple, nous pouvons caractériser la relation entre comportements individuels et comportements collectifs de la manière suivante :

- Chaque agent est sensible à un certain ensemble de paramètres de son environnement *physique* (ici la température), cette sensibilité étant une caractéristique intrinsèque de l'agent,
- chaque agent peut utiliser une information sur les variations de ces paramètres pour initier des actions qui vont influencer en retour la valeur de ces paramètres²,
- l'action conjuguée des différents sujets, médiatisée par une variable d'environnement à portée collective (ici la température), a des effets

²On parlera de *rétroaction négative* lorsque cette action va à l'encontre de cette variation, comme c'est le cas dans l'exemple présenté ici, et de *rétroaction positive* lorsque la cette action amplifie cette variation (comme c'est le cas par exemple dans la construction des termitières).

dans l'environnement qui se réalisent à un niveau macro par rapport aux sujets (ici régulation de la température globale moyenne de la ruche). Se sont souvent ces effets que l'on qualifie de *phénomènes collectifs émergents*.

Deux distinctions importantes sont à relever entre l'exemple précédent et ce qui nous semble proprement humain dans le fonctionnement de nos sociétés³ :

1. Les variables d'environnement auxquelles sont sensibles les agents dans l'exemple précédent sont des caractéristiques intrinsèques des agents. Dans les sociétés humaines, celles-ci sont également le fruit d'une évolution culturelle,
2. Dans l'exemple précédent, les variables d'environnement correspondent à des phénomènes physiques. Les autres sujets ne sont pris en compte que dans la mesure où leur présence a un effet physique sur l'environnement. En ce qui concerne les êtres humains, ce n'est pas être polémique que de dire que les individus prennent également en compte ce que sont les partenaires de leurs interactions en termes de croyances, d'intentions, de valeurs, etc.⁴ et que l'influence sociale agit également à ce niveau là de la constitution des individus.

Dans ces deux distinctions résident plusieurs grandes questions de la sociologie à savoir : Quelle est la relation entre individus et structures collectives ? Quel est la part du social dans la constitution de l'individu ? Quelle est l'influence de l'individu sur les structures collectives ? Dans quel sens et dans quelle mesure sommes-nous influencés à travers les liens sociaux que nous entretenons dans les décisions que nous prenons envers nous-mêmes et envers les autres ?

Ces questions sont étroitement liées au fait que chez l'Homme plus que chez toute autre espèce animale, les nouveau-nés sont fondamentalement incomplets à leur naissance. Toutes les histoires d'enfants sauvages ont suffisamment intrigué notre imaginaire pour attirer notre attention sur le fait que les déterminants biologiques ne jouent qu'un rôle partiel dans le futur développement du nouveau-né en tant qu'être social. Cette indétermination de l'*Être* humain a été maintes fois soulignée par Tarde et est l'une des principales remarques motivant sa théorie sociologique :

La fonction cérébrale, l'esprit, se distingue des autres fonctions en ce qu'elle n'est pas une simple adaptation à une fin précise par un moyen précis, mais une adaptation à des fins multiples et indéterminées qui doivent être précisées plus ou moins fortuitement par le moyen même qui sert à les poursuivre et qui est immense, à savoir par l'imitation [...]. [Tarde(1890)]

Ce court passage condense de manière extraordinaire la problématique que nous avons exposée plus haut. Tarde y défend l'idée que les fins guidant un individu dans sa vie sociale sont elles-mêmes le produit de cette vie en société mais ne pré-existent pas l'activité collective individus. Elles s'élaborent chemin faisant. Tarde suggère ainsi que ces fins qui conditionnent le traitement de l'information au niveau individuel sont le

³Mentionnons au passage une littérature très intéressante en éthologie comparée où les germes de ce qui nous semble proprement humain se retrouvent dans certaines sociétés animales évoluées, notamment chez certains primates et certaines espèces de cétacés. Voir par exemple les excellents ouvrages [Lestel(2001)] et [Tomasello(1999)].

⁴*cf.* [Rabin(2002)] pour des exemples d'études expérimentales sur l'influence des intentions dans l'interaction.

produit de l'enchevêtrement au niveau collectif de ces mêmes traitements de l'information. Cet enchevêtrement instancie la boucle réflexive des fins sur elles-mêmes leur permettant de s'auto-constituer. Tarde donne son origine : l'*imitation*. La dynamique que ce mode de fonctionnement engendre est précisément ce que nous avons désigné sous le terme de *cognition sociale*.

2 L'imitation : trait d'union entre l'individu et le collectif

Pour commencer par des généralités, soulignons que la notion d'imitation, contrairement à beaucoup de ses présentations dans la littérature des sciences sociales, est bien différente de la notion de conformisme. La première est un principe, la seconde est une loi de comportements, elles n'ont pas le même degré de généralité.

Par ailleurs, il ne s'agit pas ici de donner l'exclusivité à l'imitation dans l'organisation des conduites humaines ou de prendre la position extrême de Tarde suivant laquelle « *tout n'est socialement qu'inventions et imitations* » [Tarde(1890)]. Certains comportements sont innés, d'autres acquis. Les comportements acquis dépendent de processus de décision qui résultent d'un apprentissage. Pour rester dans les généralités, les processus de prise de décision chez les êtres humains dépendent de deux grandes catégories d'apprentissage : l'apprentissage individuel et l'apprentissage social. L'imitation, quelque soit la définition que l'on en adopte, recouvre une large part de l'apprentissage social. Les processus mimétiques se répartissent eux-mêmes en deux catégories : les processus automatiques et/ou inconscients qui sont présents dès la naissance chez le nouveau-né [Meltzoff and Prinz(2002)] et peuvent prendre la forme à l'âge adulte d'un certain conformisme ; les processus réfléchis, que l'on qualifie souvent d'imitation rationnelle et que Tarde [Tarde(1890), Ch. V] attribuait à des *causes logiques*. Il est entendu qu'une théorie de l'action doit prendre en compte ces différents types de prise de décision, mimétique ou non.

Dans ce qui suit, nous allons isoler les phénomènes mimétiques des autres types d'apprentissage dans le but d'identifier clairement les réponses qu'ils peuvent apporter aux questions que nous nous sommes posées. Plus particulièrement, nous nous intéresserons aux causes logiques de l'imitation.

2.1 Les approches historiques de l'imitation

L'imitation a été considérée dans beaucoup de systèmes épistémologiques comme un lieu de démarcation entre sociétés humaines et autres sociétés animales. On lit par exemple chez Aristote (*Poétique* 4) "*L'Homme diffère des autres animaux en ce qu'il est plus apte à l'imitation*". A la fin du XIXème siècle, les sciences humaines et sociales ont connu un engouement sans précédent pour l'imitation en tant que principe structurant commun à la cognition individuelle et la cognition sociale. Alors que James Marc Baldwin [Baldwin(1897)], un des pères de la psychologie expérimentale, plaçait l'imitation comme clef de voûte de sa théorie du développement et notamment de l'émergence de la volition, Gabriel Tarde promettait aux lois générales qui régissent la répétition imitative un rôle analogue en sociologie à celui de l'hérédité en biologie.

Ces deux approches se rejoignent autour de deux convictions d'une étonnante actualité, dont la portée épistémologique est de première importance pour les sciences sociales :

- Une théorie du social doit de se doter de ses propres outils conceptuels et formels. Pour cela il serait trompeur de chercher à donner aux sciences sociales un 'air biologique' ou un 'air mécanique' afin de récupérer les outils formels d'autres disciplines car ce serait "*lâcher la proie pour l'ombre*"⁵. Autrement dit, les dynamiques sociales sont fondamentalement irréductibles aux dynamiques gouvernant les systèmes naturels⁶. La raison principale de cette conviction : pour Tarde comme pour Baldwin, ces dynamiques sont par essence d'origine mimétique.
- Il y a dans les systèmes sociaux co-détermination des individus et du collectif, celle-ci étant rendue possible par l'imitation. Les systèmes sociaux engendrent ainsi de façon interne leur propre espace de significations qui est la projection au niveau collectif d'un processus dialectique intrinsèquement social de constitution du sens. Celui-ci s'initie dès l'émergence de la distinction entre soi et autrui : "*My sense of myself grows by imitation of you, and my sense of yourself grows in terms of my sense of myself*"⁷.

Le premier point est important dans une approche heuristique car il pose des contraintes sur les types de modèles, formels en particulier, qui sont susceptibles de nous éclairer sur les faits sociaux. Avec notre recul historique, cette intuition est renforcée par les tentatives infructueuses de l'économie "orthodoxe" et de la sociobiologie pour expliquer certains faits sociaux, insuccès révélés en particulier par leurs difficultés à rendre compte de la diversité des fins.

Le deuxième point est tout aussi important car c'est un programme de recherche qui se déploie entre individualisme méthodologique et holisme, préfigurant l'*individualisme méthodologique complexe* [Dupuy and Dumouchel(1983)]. Les alternatives qu'il propose peuvent constituer des apports substantiels aux débats qui font rage entre ces deux prises de position. En particulier, il y avait à l'époque de Tarde l'idée que l'imitation délibérée est le premier signe distinctif de la volition et de l'action raisonnée au cours du développement : la manière dont l'individu se démarque de son environnement social tient précisément à une modification de la forme initialement innée de l'imitation, c'est donc un espace de libre choix créé à l'intérieur même du domaine réservé aux processus de détermination sociale.

2.2 L'imitation aujourd'hui : entre holisme et individualisme méthodologique

De manière étonnante, ce programme de recherche n'a pas trouvé de successeur immédiat. Après avoir été quelque peu délaissé, le thème de l'imitation a cependant été repris par de grandes figures des sciences sociales : Piaget en psychologie, Girard [Girard(1961)] en anthropologie, Keynes et Hayek en économie [Dupuy(2004)] ; pour être aujourd'hui un sujet d'une actualité brûlante aussi bien en psychologie sociale [Bandura(1977)], qu'en anthropologie évolutionniste ([Boyd and Richerson(1985)], [Donald(1991)], [Henrich and Boyd(1998)]), en économie ([Orléan(2002)], [Selten and Apesteguia(2005)]),

⁵[Tarde(1890), Intro. au CH I]

⁶Même si des réductions partielles n'en restent pas moins instructives.

⁷[Baldwin(1897)]

en éthologie ([Byrne and Russon(1998)], [Tomasello(1999)]) ; et également en psychologie, neurologie, robotique ou philosophie de l'esprit (voir par exemple [Nadel and Decety(2002)]).

Mais force est de constater qu'en traversant les disciplines, le thème de l'imitation a perdu son originalité première, chaque théorie ne déployant qu'un aspect de l'imitation. Ceci est particulièrement visible en économie où certains chercheurs comme André Orléan ou Jean-Pierre Dupuy ont pu privilégier une approche girardienne des processus mimétiques, opaques aux individus, qui rappelle de ce que Baldwin nommait imitation simple et que Tarde rangeait dans les imitations relevant de causes 'non-logiques' ; alors que d'autres comme [Selten and Ostmann(2001)] ou [Frank(2003)] privilégient une approche rationnelle de l'imitation, dans le prolongement de la notion d'imitation persistante chez Baldwin et de l'imitation relevant de causes logiques chez Tarde.

D'un côté, nous avons l'idée qu'un individu désire parce qu'il imite, de l'autre celle qu'il imite parce qu'il désire. Formellement, cela se traduit par des approches très différentes. Pour le courant que l'on pourrait qualifier de girardien, la valeur des choses, des opportunités d'action se lisent dans les désirs des autres. L'imitation est alors en quelque sorte un conformisme englobe le niveau des désirs ou des fins, chacun recherchant ce qu'il pense deviner chez l'autre. Mais le fait qu'un individu ait telle ou telle structure sur ses désirs ne change pas la manière dont il subira cette influence mimétique. La manière dont un individu choisit ses modèles, la forme de l'imitation, est une propriété commune à tous les individus.

En revanche, pour le courant qui s'attache à l'aspect logique et rationnel de l'imitation, celle-ci est au service de principes ou de fins pré-établis : A imite B (en adoptant sa manière d'agir ou sa stratégie) s'il estime que ce dernier est particulièrement performant suivant des critères reconnus par A , ces critères étant une propriété permanente de A (par exemple, si A est un individu cupide, il aura certainement une propension à imiter particulièrement des individus qui se seront fait remarquer par leur opulence). La forme de l'imitation est alors une caractéristique individuelle.

Outre le fait que ces approches partielles de la notion d'imitation ont pour conséquences des débats interminables sur la définition même d'imitation, le fait de ne considérer qu'une seule face de la médaille rompt le lien que fait l'imitation entre cognition individuelle et cognition sociale et fait disparaître du même coup cette possibilité de voie médiane entre individualisme méthodologique et holisme.

2.3 L'imitation, trait d'union entre l'individu et le collectif : les jeux metamimétiques

L'étude de l'imitation semble donc ouvrir une voie vers cette "*dialectique réursive*" par laquelle "*la pensée des moyens de la fin [devient] le moyen de repenser les fins*" [Le Moigne(1988)]. Pour autant, le cadre conceptuel posé par Tarde reste flou sur les moyens à mettre en oeuvre en pratique pour saisir par des modèles de pensée les propriétés qualitatives qu'apportent ce mode d'auto-transformation du social. Les arguments qu'il utilise sont essentiellement des descriptions de phénomènes. Il leur manque comme complémentaire des reconstructions de faits stylisés ou des ébauches de modèles qui nous permettrait de penser plus précisément la relation entre individus et collectif. Tarde nous donne cependant un indice dans la définition qu'il donne des causes logiques de l'imitation [Tarde(1890)] :

« Les causes sociales sont de deux sortes : logiques ou non logiques. ... Les causes logiques agissent quand l'innovation choisie par un homme l'est parce qu'elle est jugée par lui plus utile ou plus vraie que les autres, c'est-à-dire plus d'accord que celles-ci avec les buts ou les principes déjà établis en lui (par imitation toujours). »

La quasi-totalité de cette citation correspond à peu de choses près à la description que l'on peut attendre de l'imitation rationnelle de la part d'un économiste : un individu choisit une action plutôt qu'une autre, choisit d'imiter untel plutôt que tel autre, parce que le résultat attendu de cette imitation correspond à la meilleure évaluation sur l'échelle de ses préférences. Mais Tarde rajoute, sans cependant s'étendre, « *par imitation toujours* », suggérant ainsi que les buts et les principes des individus eux-mêmes évoluent suivant une dynamique mimétique.

Mais dans ce cas, quels sont les buts et les principes qui servent de référence lorsqu'il s'agit de changer par imitation ses buts et ses principes ? Tarde ne semble pas suggérer que cette seconde imitation au niveau meta soit de nature différente que celle qui agit au niveau des actions (par exemple qu'elle relèverait de causes non logiques, ce qui nous rapprocherait de Girard). Ce qu'il propose plutôt c'est que l'on doit dissocier le *principe d'imitation* des *critères* guidant les actes d'imitation, ces derniers pouvant également faire l'objet d'un acte d'imitation.

Cette approche trouve une illustration dans le formalisme des jeux métamimétiques [Chavalarias(2004)]. Celui-ci s'appuie sur les relations possibles entre la spécificité de l'imitation humaine et certaines capacités cognitives propres aux être humains telles que la réflexivité ou la méta-cognition. En particulier, cela permet de formaliser cette dissociation entre principe d'imitation et critères d'imitation. L'imitation d'un nouveau moyen pour satisfaire une ancienne fin devient alors le principal facteur d'adoption de nouvelles fins.

D'un point de vue formel, on montre ainsi que les règles d'imitation peuvent être leurs propres métarègles : une règle d'imitation, en tant que processus permettant de sélectionner un modèle, puis un trait à copier, peut être utilisée pour choisir parmi les modèles possibles, celui auquel un agent empruntera sa nouvelle règle d'imitation. Cette équivalence d'un niveau et de son meta niveau « quelque part » dans la description de la dynamique rejoint les caractéristiques fondamentales des systèmes capables d'auto-transformation telles qu'elles ont pu être décrites par des auteurs comme [Hofstadter(1999)].

Déclinée à travers le prisme de l'hétérogénéité des agents, cette représentation formelle de l'imitation permet de rendre compte d'une auto-organisation des fins au niveau d'une population au sens de sélection spontanée d'une distribution de fins particulières à partir de l'ensemble des fins imaginables par les agents. Les configurations sociales stables dans ce cadre correspondent à un nouveau concept d'équilibre, les *états contrefactuellement stables* : « aucun individu ne peut s'imaginer mieux qu'il n'est en se mettant contrefactuellement à la place d'autrui ». Le chemin que parcourt un système pour arriver à ce type d'équilibre (ou d'attracteurs le cas échéant) correspond précisément à ce que nous avons appelé plus haut un processus de *cognition sociale*.

Cette approche des dynamiques sociales se révèle en rupture avec les approches formelles traditionnelles qui se placent généralement dans un cadre finaliste où les agents optimisent une quantité qui pré-existe à leur activité. Elle interprète en effet l'hétérogénéité auto-organisée des sys-

tèmes sociaux humains comme une différenciation par co-évolution d'une multiplicité de critères possibles, qui dès lors ne peuvent se concevoir que de manière contextuelle et d'un point de vue individuel. Nous renvoyons le lecteur aux travaux [Chavalarias(2006)] pour un exposé plus détaillé du cadre formel des jeux métamimétiques et [Chavalarias(forthcoming 2007)] pour un exemple d'application à l'émergence de la coopération. Nous allons maintenant fournir un exemple minimal pour montrer en quoi ce type d'approche peut nous aider à penser l'auto-transformation du social en termes de cognition sociale.

3 Une représentation de la cognition sociale autour d'un jeu metamimétique

3.1 Comment représenter l'auto-transformation des fins

De même que le cercle griffonné sur une feuille de papier permet de penser les propriétés du cercle bien qu'il n'en ait quasiment aucune, nous allons maintenant exposer les grandes lignes d'un modèle multi-agents susceptible d'illustrer les propriétés d'auto-transformation et les relations individus-collectif que nous avons ébauchées. Soulignons qu'il ne s'agit pas de dresser un portrait réaliste des systèmes sociaux mais de s'intéresser à des types de systèmes tels que les relations entre les différents niveaux d'organisation qu'ils mettent en scène nous permettent de réfléchir sur les relations entre les différents niveaux d'organisation composant les systèmes sociaux.

Revenons sur cette "*adaptation à des fins multiples et indéterminées qui doivent être précisées plus ou moins fortuitement par le moyen même qui sert à les poursuivre [...] à savoir l'imitation.*" Si nous prenons Tarde à la lettre, le modèle le plus simple que nous pouvons imaginer des transformations des fins est de supposer qu'une fin est modifiée au cours d'un processus d'imitation où elle intervient elle-même comme critère dans l'acte d'imitation. Elle est alors à la fois l'objet de la transformation et son motif. Par exemple, il se peut très bien qu'afin de pouvoir faire de la recherche sur vos thèmes de prédilection (fin1), il vous semble incontournable de recueillir vous-même des financements (fin2), cette seconde fin finissant par occuper en tâches administratives tout le temps initialement consacré à la recherche.

Ce type de glissement est fort fréquent. Nous allons maintenant réfléchir sur l'une de ses schématisations dont la simplicité aura l'avantage de nous permettre d'appréhender les phénomènes collectifs. Ce glissement peut se caricaturer de la manière suivante : « si au regard de la fin 1 que je poursuis je m'aperçois que poursuivre la fin 2 est l'une des meilleures solutions pour arriver à mes fins, alors je vais consacrer du temps à poursuivre cette seconde fin. » Ici, le verbe 'apercevoir' rend compte d'un apprentissage. Dans les limites que nous nous sommes fixées, celui-ci est d'ordre mimétique. Ceci nous permet de réinterpréter la phrase de la manière suivante : « si en me mettant contrefactuellement à la place de l'une de mes connaissances, je me trouve dans une situation meilleure relativement aux fins que je poursuis actuellement, je vais chercher à m'inspirer du comportement et/ou des fins de cette personne ». La question est maintenant de savoir ce qu'il se passe lorsque toute une population raisonne de la sorte. Est-ce que certains individus vont se stabiliser sur des fins particulières ?

TAB. 1 – Matrice donnant la structure de l’interaction sociale stylisée en terme de gains matériels

Joueur <i>A</i>	Joueur <i>B</i>	
	<i>C</i>	<i>D</i>
<i>C</i>	(R, R)	(S, T)
<i>D</i>	(T, S)	(P, P)

Va-t-on assister à l’émergence de groupes partageant des fins similaires ? Quel type de dynamiques obtient-on sur les fins et quels types de relations individus/collectif cela implique-t-il ?

3.2 Une représentation stylisée d’une situation d’interaction sociale

Une des manières de répondre à ces questions est de considérer effectivement un ensemble d’agents raisonnant de la sorte, capables imaginer et de suivre un ensemble circonscrit de fins et d’actions, puis d’en étudier les dynamiques collectives. Il nous faut pour cela imaginer une situation stylisée d’interaction sociale : un cadre d’interaction et l’ensemble des représentations dont peuvent s’en faire les agents.

Nous nous placerons pour cet exemple dans le cadre de la théorie des jeux et nous prendrons le modèle d’interaction le plus simple et le plus répandu dans ce domaine. Celui-ci consiste en une interaction dyadique entre deux joueurs qui se répète dans le temps. A chaque période de temps, les joueurs ont à choisir entre deux comportements possibles à opposer à leur partenaire, que nous désignerons par *C* et *D*. Nous considérerons un cas particulier où l’interaction se traduit pour chacun des joueurs par des gains ou des pertes matérielles (le plus classique étant d’imaginer des gains monétaires) :

- si chacun des agents choisit *C*, le gain individuel est de *R*,
- si chacun des agents choisit *D*, le gain individuel est de *P*,
- s’ils choisissent des actions différentes, celui qui a choisi *D* gagne *T* et celui qui a choisi *C* gagne *S*.

Cette structure de jeu peut se résumer par une matrice d’allure bien connue des théoriciens des jeux (Table1).

L’étude que nous présentons ici porte sur un jeu célèbre appelé dilemme du prisonnier qui traduit le type de conflit qu’il peut y avoir entre intérêts individuels et intérêt collectif. La seule caractéristique dont nous aurons besoin cependant dans cette brève analyse⁸ est le fait que, quelque soit l’action de son partenaire, un joueur a des gains supérieurs s’il joue *D* que s’il joue *C*. Une fois défini le type d’interaction, il nous reste à définir

la structure du réseau social que nous allons étudier et les représentations que peuvent se faire les agents de ces interactions.

Nous placerons les agents sur une grille carrée torique⁹, ceux-ci ayant la possibilité d’interagir avec leurs huit voisins (fig.1).

En ce qui concerne les représentations que se font les agents de leurs interactions celles-ci vont dépendre, comme nous l’avons déjà mentionné,

⁸Voir [Chavalarias(forthcoming 2007)] pour une étude plus détaillée.

⁹Imaginez un échiquier dont les pièces sont les agents et dont les bords opposés se correspondent.



FIG. 1 – Un agent avec ses huit voisins.

des catégories perceptives dont disposent les agents et de la manière dont ils les agencent pour se représenter leurs fins. Nous ne nous intéresserons pas ici à la manière dont se forment ces catégories perceptives, qui est un autre problème d'évolution culturelle, mais uniquement à la manière dont s'auto-organisent les fins possibles. Nous ferons donc les hypothèses suivantes :

- les agents peuvent percevoir les gains des partenaires avec ils interagissent,
- les agents peuvent percevoir le comportement et les fins de leurs partenaires,

En ce qui concerne les opérations que peuvent faire les agents sur ces perceptions élémentaires nous supposons que :

- les agents peuvent calculer les densités de comportements et de fins parmi leurs partenaires,
- les agents peuvent calculer le maximum de deux quantités,
- les agents peuvent calculer le minimum de deux quantités.

Cela nous donne quatre fins possibles :

- être l'agent dont les gains sont les plus élevés dans le voisinage,
- être l'agent dont les gains sont les moins élevés dans le voisinage,
- être dans la majorité relativement aux comportements et aux fins,
- être dans la minorité relativement aux comportements et aux fins.

Traduit en terme de règles d'imitation, cela nous donne quatre règles possibles souvent utilisées dans la littérature (table 2) :

1. *Maxi* : « copie l'agent de ton voisinage dont les gains sont plus élevés »,
2. *Mini* : « copie l'agent de ton voisinage dont les gains sont les moins élevés »,
3. *Conformisme* : « copie la règle ou le comportement majoritaire parmi tes voisins »,
4. *Anti-conformisme* : « copie la règle ou le comportement minoritaire parmi tes voisins ».

Nous attirons l'attention sur le fait que l'agent étant dans son propre voisinage, il ne s'engagera pas dans un acte d'imitation s'il est effectivement l'un des « meilleurs » agents de son voisinage. On dira alors qu'il est *contrefactuellement stable*.

TAB. 2 – L'espace des fins possibles traduites en terme de règles d'imitation.

Opérations	Perceptions	
	Densités	Gains
<i>max</i>	<i>Conformisme</i>	<i>Maxi</i>
<i>min</i>	<i>Anti-conformisme</i>	<i>Mini</i>

3.3 La cognition sociale : quand le changement va changeant

Autant nous pouvons imaginer ce qu'il se passe lorsque deux agents jouent à ce jeu metamimétique, autant cela dépasse nos capacités de représentation lorsque la population comporte des milliers d'agents. Si nous voulons étudier les relations individus/collectif, il nous faut passer par un médium externe ou une représentation analytique. Nous avons choisis de recourir aux simulations multi-agents et nous présentons ici l'effet de comportements metamimétiques à l'échelle d'une population de 10 000 agents en réalisant *in silico* la situation stylisée d'interaction sociale que nous venons de décrire (cf. figure 2).

Plusieurs choses sont frappantes aux premiers abords lorsque l'on étudie l'évolution d'un tel système artificiel et d'étroits parallèles peuvent être tracés avec la pensée de Tarde.

"La logique de l'imitation conduit à ce que chaque homme se met finalement d'accord avec lui-même par l'adoption d'un système logique d'idées et d'une conduite conséquente." [Tarde(1890)]

Tout d'abord, après une courte période d'intense réorganisation pendant laquelle la plupart des agents changent leur règle et leur comportement, la population finit par se stabiliser dans un état quasi statique (attracteur) où seulement quelques agents continuent de changer périodiquement leur règle ou leur comportement. Dans cet état, la plupart des agents sont contrefactuellement stables : ils sont en équilibre avec leur environnement par rapport aux règles qu'ils ont adoptés.

"De là cette définition du groupe social : une collection d'êtres en tant qu'ils sont en train de s'imiter entre eux ou en tant que, sans s'imiter actuellement, ils se ressemblent et que leurs traits communs sont des copies anciennes d'un même modèle."

Observons ensuite que cet état collectif est fondamentalement *hétérogène*. Contrairement à ce que l'on pourrait penser, l'imitation lorsqu'elle s'exerce au niveau meta n'aboutit pas à une uniformisation mais bien à une différenciation. Cet état collectif est organisé en sous-groupes d'agents dont les structures sont caractéristiques des types d'agents qui les composent. Ces structures nous permettent même de deviner le code couleur de la figure 2 : les conformistes sont structurés en larges zones plus ou moins convexes à l'intérieur desquelles ils sont effectivement en majorité. À l'intérieur de ces zones, leur comportement est uniforme mais peut être indifféremment *D* ou *C*. Les anti-conformistes sont éparpillés sur tout le territoire, ils sont isolés et globalement en minorité. Leur comportement est en général contraire à ce qui est pratiqué dans leur voisinage. Les populations d'agents maxi et mini quant à elles sont la plupart du temps enchevêtrées, les maxi ayant adopté l'action *D* associée aux gains les plus

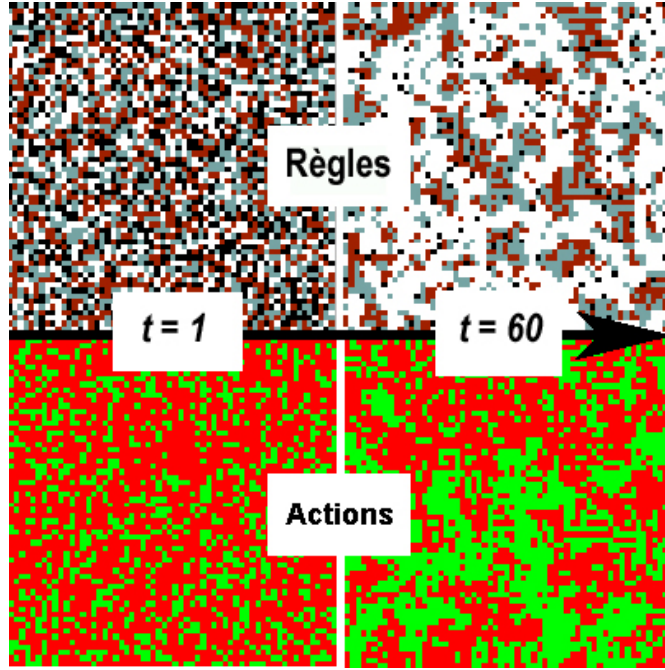


FIG. 2 – Evolution de la distribution spatiale des règles d'imitation (partie supérieure) et des comportements (partie inférieure) à partir d'un état initial désordonné vers un attracteur structuré. Chaque petit carré représente un agent. Chaque agent interagit avec ses huit voisins comme sur la figure 1. La configuration à l'attracteur est globalement stable à l'exception de quelques agents qui ne cessent de changer leur règle ou leur comportement. *Code couleur* : pour les règles - blanc : les conformistes, noir : les anti-conformistes, gris clair : mini, gris foncé : maxi. Pour les comportements - gris clair : action C , gris foncé : action D . La population comporte ici 10 000 agents. Dans l'état initial de gauche, les règles sont assignées aux agents de manière aléatoire uniforme, les comportements sont assignés de manière aléatoire en respectant une proportion initiale de 30% de comportements C . L'état de droite correspond à l'état de la population après 60 périodes de jeu.

forts, et les mini ayant adopté l'action C associée au gain les plus faibles : les uns profitent des autres qui s'en réjouissent.

Ces « groupes sociaux artificiels » sont la projection au niveau collectif de l'espace des représentations envisageables par les agents, que Tarde désignait comme l'ensemble des « *virtualités élémentaires* ». Leur formation en tant que macro-structures est une conséquence des interactions mimétiques répétées entre agents au cours desquelles des représentations potentielles s'actualisent dans un contexte social.

"Quelque chose de bien plus important qu'une simple augmentation de différence s'accomplit dans l'univers incessamment, la différenciation de la différence elle-même. Le changement même y va changeant, et dans un certain sens qui, d'une ère de différences crues et juxtaposées, comme de couleurs criardes et non fondues, nous achemine à une ère de différences harmonieusement nuancées."

Comme nous l'avons défendu, l'imitation telle que nous l'avons présentée doit être envisagée dans les modèles comme un principe et non comme une loi : nous n'obtenons une loi ou une règle d'imitation qu'à partir du moment où l'on définit le critère par rapport auquel doit s'effectuer l'imitation¹⁰. Cela signifie que la loi d'évolution globale du système d'agents est en fait définie par l'arrangement spatial des règles d'imitation des différents agents. La partie supérieure de la figure 2 est en quelque sorte un dessin de cette loi d'évolution à deux périodes différentes. La différence du dessin entre ces deux périodes signifie donc que cette loi d'évolution elle-même va changeant au cours de l'évolution de la société artificielle. C'est en ce sens que l'on peut dire qu'il y a auto-transformation du système au cours de son fonctionnement. Cette dynamique qui agit autant au niveau des comportements qu'au niveau des règles d'évolution de ces comportements est une illustration de ce que nous avons appelé *cognition sociale*.

3.4 Holisme ou individualisme méthodologique ?

Si maintenant nous considérons cet exemple à la lumière des deux grands paradigmes des sciences sociales, individualisme méthodologique et holisme, nous pouvons constater qu'il ne relève ni de l'un, ni de l'autre mais d'un troisième type, l'*individualisme méthodologique complexe*. Dire que l'évolution de cette société artificielle est le produit des interactions d'individus déjà constitués hors de la société n'a aucun sens car nous voyons ici que l'identité des individus change au cours du jeu social et que leurs téléologies individuelles se constituent précisément au cours de leur évolution dans un contexte social.

Mais dire que l'évolution de cette société est dirigée par des régularités qui préexistent aux interactions des agents n'a pas non plus de sens. Certes, nous pourrions objecter que l'espace des représentations que nous avons considéré est ici donné en dehors des agents, qu'il pré-existe à leur activité. Rappelons simplement que nous avons laissé de côté la constitution de cet espace pour des raisons de clarté de l'exposé. Comme le souligne Tarde dans son oeuvre, cet espace est en fait le produit de

¹⁰Nous attirons l'attention du lecteur sur le fait que le titre du premier ouvrage de Tarde est bien « Les lois de l'Imitation » et non « La loi de l'Imitation » ce qui souligne l'idée d'un principe duquel peuvent découler plusieurs types de régularités.

l'évolution culturelle. C'est un espace ouvert où certaines innovations individuelles sans cesse apparaissent et ne perdurent dans le contexte social que dans la mesure où elles sont reprises et amplifiées par la dynamique mimétique collective. L'adoption de nouveaux principes, de nouvelles valeurs, de nouveaux traits sociaux, comme on peut le voir par exemple dans l'évolution d'une langue, sont conditionnées par la meta-dynamique définie par l'arrangement des entités sociales existantes, mais chaque individu porte en lui la possibilité d'introduire l'innovation qui un jour à terme modifiera de manière significative cette méta-dynamique.

Enfin, soulignons que l'interprétation de cette dynamique en termes d'optimisation globale ou de finalisme n'a aucun sens. Allez dire à un anticonformiste qu'il a perdu parce qu'il s'est retrouvé en minorité et il vous rira au nez.

4 Conclusion

Nous avons souhaité dans cet article à préciser le concept de *cognition sociale* en insistant sur son caractère spécifique aux sociétés humaines en tant que processus qui aboutit à l'auto-organisation des téléologies individuelles.

Nous avons ainsi vu l'importance de considérer dans une approche des systèmes sociaux non seulement les dynamiques sur les comportements mais également les meta-dynamiques sur les règles / croyances / représentations des agents qui déterminent ces comportements. Ces meta-dynamiques sont en effet à même d'exprimer la particularité d'auto-transformation du social.

Le terme cognition est alors envisagée sous un angle qui nous rappelle la notion de *clôture opérationnelle* introduite par Francisco Varela [Varela(1983)]. Il s'agit d'un traitement de l'*in*-formation : les sociétés forment leurs propres constituants au cours de leur fonctionnement.

Comme le défendait Tarde en d'autres termes, nous avons vu que l'imitation était effectivement susceptible de rendre compte de manière très concrète de ce phénomène de cognition sociale.

Dès lors, celle-ci exprime une différenciation sociale par co-évolution culturelle, la culture pouvant se définir comme *l'individuation par le collectif* [Lestel(2001)]. Elle aboutit à l'émergence de groupes sociaux et de normes, ceux-ci apparaissant à un observateur extérieur comme des macro-régularités distinctes de la somme des activités individuelles. Un exemple simple nous a permis de montrer en quoi une approche système complexe est susceptible de nous aider à envisager ce type de processus et à appréhender "l'inséparable conjonction d'un fonctionnement et d'une transformation dans un contexte et par rapport à quelques projets eux-mêmes changeants." [Le Moigne(1988)]. L'essentiel des avancées théoriques est encore devant nous.

Remerciements

Cet article est le produit d'influences trop nombreuses pour être mentionnées individuellement. Qu'elles en soient collectivement remerciées. Je remercie en particulier l'ensemble des membre du CREA pour leurs discussions toujours stimulantes, ainsi que Pascal Roggero pour son invitation dans ce réseau de l'AFS. Cet article a bénéficié du soutien matériel de l'École Polytechnique.

Références

- [Baldwin(1897)] Baldwin, J.M., 1897. Le développement mental chez l'enfant et dans la race. F. Alcan.
- [Bandura(1977)] Bandura, A., 1977. Social Learning Theory. Prentice-Hall, Englewood Cliffs, N.J.
- [Bonabeau and Theraulaz(1994)] Bonabeau, E., Theraulaz, G., 1994. Intelligence Collective. Hermès, Paris.
- [Bourgine(2004)] Bourguine, P., 2004. What is cognitive economics? In : P. Bourguine, J.P. Nadal (Eds.), Cognitive Economics – An interdisciplinary approach, Berlin : Springer. pp. 1–12.
- [Boyd and Richerson(1985)] Boyd, R., Richerson, P. J., 1985. Culture and the evolutionary process. Chicago : Chicago University Press.
- [Byrne and Russon(1998)] Byrne, R. W., Russon, A. E., 1998. Learning by Imitation : a Hierarchical Approach. Behavioral and Brain Sciences .
- [Chavalarias(2004)] Chavalarias, D., 2004. Métadynamiques en Cognition Sociale. Ph.D. thesis, Ecole Polytechnique, Paris, France.
- [Chavalarias(2006)] Chavalarias, D., 2006. Metamimetic games : Modeling Metadynamics in Social Cognition. Journal of Artificial Societies and Social Simulations Vol. 9, Issue 2.
- [Chavalarias(forthcoming 2007)] Chavalarias, D., forthcoming 2007. Cooperation as the outcome of a social differentiation process in metamimetic games. In : E. Bruce, C. Hernández Iglesias, K.G. Troitzsch (Eds.), Social Simulation : Technologies, Advances and New Discoveries.
- [Conein(2005)] Conein, B., 2005. Les sens sociaux. Economica.
- [Donald(1991)] Donald, M., 1991. Origins of the Modern Mind. Cambridge, MA : Harvard University Press.
- [Dupuy(2004)] Dupuy, J.P., 2004. Economic as Symptom. In : Transforming Economics - Perspectives on the Critical Realist Project, London and New York : Routledge. pp. 227–251.
- [Dupuy and Dumouchel(1983)] Dupuy, J.P., Dumouchel, P. (Eds.), 1983. L'auto-organisation, de la physique au politique. Seuil.
- [Frank(2003)] Frank, J., 2003. Natural selection, rational economic behavior, and alternative outcomes of the evolutionary process. Journal of Socio-Economics 32, 601–622.
- [Girard(1961)] Girard, R., 1961. Mensonge romantique et Vérité romanesque. Grasset.
- [Henrich and Boyd(1998)] Henrich, J., Boyd, R., 1998. The Evolution of Conformist Transmission and the Emergence of Between-Group Differences. Evolution and Human Behavior 19, 215–241.
- [Hofstadter(1999)] Hofstadter, D.R., 1999. Gödel, Escher, Bach : An Eternal Golden Braid. Basic Books.
- [Hutchins(1995)] Hutchins, E., 1995. Cognition in the wild. MIT Press : Cambridge, MA.
- [Le Moigne(1988)] Le Moigne, J.L., 1988. Représenter et Raisonner les Comportements Socio-Economiques. In : C. Roland-Levy, P. Adair (Eds.), Psychologie économique, théorie et applications, Economica. pp. 317–337.

- [Lestel(2001)] Lestel, D., 2001. Les origines animales de la culture. Flammarion.
- [Meltzoff and Prinz(2002)] Meltzoff, A.N., Prinz, W. (Eds.), 2002. The Imitative Mind : Development, Evolution, and Brain Bases. Cambridge University Press.
- [Nadel and Decety(2002)] Nadel, J., Decety, J. (Eds.), 2002. Imiter pour découvrir l'humain, Psychologie, neurologie, robotique et philosophie de l'esprit. Puf, psychologie et science de la pensée edn.
- [Orléan(2002)] Orléan, A., 2002. Les Interactions mimétiques. In : J. Lesourne, A. Orléan, B. Walliser (Eds.), Leçons de microéconomie évolutionniste, Odile Jacob. pp. 119–157.
- [Rabin(2002)] Rabin, M., 2002. A perspective on psychology and Economics. European Economic Review 46, 657–685. A utiliser dans le débat sur les règles ad hoc.
- [Selten and Apesteguia(2005)] Selten, R., Apesteguia, J., 2005. Experimentally observed imitation and cooperation in price competition on the circle. Games and Economic Behavior 51, 171–192.
- [Selten and Ostmann(2001)] Selten, R., Ostmann, A., 2001. Imitation Equilibrium. Homo oeconomicus 43, 111–149.
- [Tarde(1890)] Tarde, G., 1890. Les lois de l'Imitation. Jean-Marie Tremblay, Les classiques des Sciences Sociales edn.
- [Tomasello(1999)] Tomasello, M., 1999. The Cultural Origins of Human Cognition. Harvard University Press.
- [Varela(1983)] Varela, F., 1983. L'auto-organisation : de l'apparence au mécanisme. In : J.P. Dupuy, P. Dumouchel (Eds.), L'auto organisation, de la physique au politique, Seuil. pp. 147–164.